Visual influence on haptic torque perception

Yangqing Xu, Shélan O’Keefe, Satoru Suzuki, Steven L Franconeri
Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208, USA; e-mail: xuy@u.northwestern.edu
Received 2 August 2011, in revised form 12 June 2012

Abstract. The brain receives input from multiple sensory modalities simultaneously, yet we experience the outside world as a single integrated percept. This integration process must overcome instances where perceptual information conflicts across sensory modalities. Under such conflicts, the relative weighting of information from each modality typically depends on the given task. For conflicts between visual and haptic modalities, visual information has been shown to influence haptic judgments of object identity, spatial features (e.g., location, size), texture, and heaviness. Here we test a novel instance of haptic–visual conflict in the perception of torque. We asked participants to hold a left–right unbalanced object while viewing a potentially left–right mirror-reversed image of the object. Despite the intuition that the more proximal haptic information should dominate the perception of torque, we find that visual information exerts substantial influences on torque perception even when participants know that visual information is unreliable.

Keywords: sensory integration, crossmodal perception, visual, haptic, weight distribution, torque perception

1 Introduction
We experience the outside world with a single integrated percept, despite the fact that sensory input from multiple modalities can conflict. Under such conflict, the relative weighting of information from each modality depends on the demands of the observer’s task. When visual information conflicts with haptic information, visual information typically dominates for tasks related to object identities and spatial features. Stronger weighting of visual information has been demonstrated in judgments of an object’s curvature (Gibson 1933), size (Rock and Victor 1964), length (Teghtsoonian and Teghtsoonian 1970), location (Hay et al. 1965; Pick et al. 1969), depth (Ho et al. 2009; Singer and Day 1969), and movement patterns (Klein and Posner 1974). In contrast, haptic information may dominate in other cases, including perception of surface texture. For example, while judging surface roughness, visual assessments were modulated by incongruent tactile information, while incongruent visual information had no effect on tactile assessments (Guest and Spence 2003). However, others have shown that observers equally weight conflicting visual and haptic information while judging surface texture (Lederman and Abbott 1981; see also Heller 1982; Lederman et al. 1986). These examples illustrate the fact that, when visual and haptic information conflicts, the relative weighting of information from the conflicting sensory modalities depends on the demands of the behavioral task.

Here we examine a novel task of torque judgments. Torque is an unbalanced distribution of forces that tend to rotate an object, which is felt as a haptic sensation. While holding an object, haptic processing of pressure distribution on the hands provides information about the imbalance in weight distribution that a grasper needs to compensate to keep holding an object. Previous research showed that, when participants grasped a visually occluded linear object at a non-balancing location and maintained it in a static orientation, they were able to reliably judge the length of the object (e.g., Burton and Turvey 1990; Carello et al. 1992; Lederman et al. 1996). This suggests that haptic modality alone can provide reliable
information about the distribution of weight along the object’s length. Nevertheless, it is possible that torque judgments may be influenced by visual information, as haptic judgments of weight are influenced by visual information. For example, visual information about volume (e.g., Amazeen 1997; Ellis and Lederman 1993; Murray et al. 1999) as well as visual information about the rotational kinematics of wielded objects (Streit et al. 2007a, 2007b) influenced haptic judgments of weight.

However, it remains unclear how visual information contributes to judgments of weight distribution (i.e., torque). In everyday life, an object’s weight distribution and its shape tend to be correlated. For example, when looking at an object that is larger on the right side than on the left side, we expect a clockwise torque when we pick up the object at the middle. Through experience, people may have learned to use this type of correlation as cues to visually estimate weight imbalance based on shape imbalance. Visually estimating weight distribution is useful because we tend to look at objects before we grasp them. Visual information thus allows us to anticipate the torque that will arise when grabbing the object off its center of gravity. Thus, in situations where visual information conflicts with haptic information regarding the direction of torque, it is possible that visual information of imbalance may be prioritized as a default despite the fact that haptic information alone is sufficient to accurately perceive torque.

Indeed, we demonstrate a powerful effect of visual information on haptic judgments of torque. In experiment 1, participants were asked to hold a left–right asymmetrically weighted object while veridical or conflicting (left–right mirror-reversed) visual information was presented unpredictably across trials. Despite the fact that participants were instructed to make haptic judgments about the direction of weight imbalance, and despite the fact that they were told that the haptic modality always provided the correct information and the visual information could be misleading, torque judgments were still strongly influenced by the visual information.

In experiment 2, participants were allowed to move the object by tilting it from side to side while making torque judgments. We predicted that this would make the task trivially easy because the consistency between participants’ intended action and the visual display would reveal the correct response. For example, if a participant purposely moved his/her hand in a clockwise direction, and the object tilted in the expected direction, then the participant would know that the monitor was displaying the veridical image, and could make the response according to the visual display. If the object tilted in the opposite direction, then the participant would know that the monitor was displaying the mirrored image, and could make the response opposite to that indicated by the visual display. We were surprised to find that the visual information still substantially influenced torque judgments, even on the mirrored trials where moving the object should have clearly revealed that the visual information was left–right mirror-reversed, demonstrating a surprisingly persistent influence of visual information on haptic torque judgments.

2 Experiment 1

Participants were asked to judge the direction of the torque of an asymmetrically weighted object while statically holding the object in its upright position. Either veridical or mirrored (i.e., conflicting) visual information was randomly and equiprobably presented on each trial.

2.1 Method

2.1.1 Participants. All twelve participants (nine females, all right-handed) had normal or corrected-to-normal vision, gave informed consent, and were given course credit for participation.
2.1.2 Stimuli and apparatus. The asymmetrically balanced object (figure 1a) consisted of a double-pronged fork with a wooden handle (34 g; 31.5 cm; center of mass 14 cm from the base), a thin wooden dowel (1 g, 22.3 cm) attached to the tines of the fork at a right angle, and a cardboard bucket (34 g, 10.16 cm in diameter) at the end of the dowel. By rotating the object, the bucket could be located either to the left or to the right of the wooden handle (and the participant’s hand). Weights (0, 113.4, or 226.8 g) were placed inside the bucket and were invisible to the participant. These weights generated torques of 0.02, 0.09, and 0.15 Nm.

The experimental apparatus included the unbalanced object, an LCD monitor (Acer V183HL 18.5 inch), and a video camera (Rocketfish HD Webcam, Model RF-HDWEB) (figures 1b and 1c). Monitor viewing distance was 28 cm. Participants held the object placed behind the monitor with their right hand by inserting their arm underneath the monitor, just to the right of the monitor stand. The video camera mounted to the rear wall of the cubicle provided a continuous live image of the object. Participants received visual information about the object only from the camera’s view projected to the monitor (the participant’s hand holding the object was invisible). The display window was located on the lower right quadrant of the monitor (figure 1c). The size and position of the object in the image were adjusted to match those of the real object, as if the display window were a rectangular hole through which participants viewed the object that they held behind the occluder. This view could either be veridical or left–right mirror-reversed, with the object’s vertical axis at the center of the display window.

Figure 1. [In color online, see http://dx.doi.org/10.1068/p7090] Experiment setup. (a) The participant held an unbalanced object with a basket of weight placed on its left or right side. While holding the stick upright, the participant haptically felt the torque in the direction of the weight. The grayed-out region was invisible to the participant. (b) A view behind the occluder shows camera placement and the experimenter holding the object until the participant grasped it. (c) The participant watched a screen showing a veridical or left–right mirrored image of the object. The image simulated how the object would appear if the screen were a rectangular window.

Participants were instructed to hold the handle and judge whether they felt the weight pulling their hand in the leftward or rightward direction (ie whether they felt a leftward or rightward torque) without allowing the object to move. Haptic information (ie torque) was manipulated across trials by alternating the side of weight bucket (left or right), and the amount of weight in the bucket. Visual information was manipulated across three conditions. The eyes-closed condition provided only haptic information. The veridical-image condition simulated a transparent window showing the actual orientation of the object, and the mirror-image condition showed the weight bucket on the opposite side.
2.1.3 Procedure. Participants put their right arm under the monitor to grab the upright object placed behind it (figure 1c). The experimenter slowly released the object, so that the participant felt the torque generated by the weight imbalance. Participants kept the object still and reported the direction of the torque that they felt (leftward or rightward by verbal reports, and pressing the corresponding buttons) as quickly as possible. Motion was not allowed in this experiment because a participant might employ a ‘trick’ to discover the correct answer based on whether the visual image moved in the same or opposite direction relative to the intended hand motion (see Heller 1992, for a similar effect). Participants were explicitly told that the visual display could show a misleading mirror image and that they should make their responses based on whether they felt their hand being pulled to the left or right. If the participant failed to keep the object still, the trial was terminated and recycled. Weight location (left or right), torque (three levels), and viewing condition (eyes-closed, veridical-image, or mirror-image) were fully crossed across 72 randomly ordered trials, with 4 trials per condition. The experiment lasted approximately 60 min.

2.2 Results and discussion

Accuracy data for experiment 1 (figure 2a) were submitted to a repeated-measures ANOVA (degrees of freedom, Greenhouse–Geisser corrected when sphericity was violated), with weight and viewing condition as the within-participant factors.

![Figure 2](image)

There was a significant main effect of viewing condition ($F_{1,14,12.59} = 8.47, p = 0.01, \eta^2 = 0.44$). Accuracy in the mirror-image condition ($M = 75.7\%$) was significantly lower than in the veridical-image condition ($M = 92.7\%$) and the eyes-closed condition ($M = 92.0\%$) ($t_{11} > 2.89, ps < 0.015$). There was no significant difference between the veridical and eyes-closed conditions ($t_{11} = 0.38, p = 0.71, ns$). There was also a significant main effect of haptic torque ($F_{1.15,12.6} = 29.74, p < 0.001, \eta^2 = 0.73$), with best performance with the strongest torque ($M = 94.8\%$), followed by the intermediate ($M = 91\%$), and weakest torque ($M = 74.7\%$) (all $t_{11} > 3.52, ps < 0.005$).

There was a significant viewing-condition-by-torque interaction ($F_{2.39,26.30} = 5.81, p = 0.006, \eta^2 = 0.35$). For the weakest torque, the accuracy in the mirror-image condition ($M = 52.08\%$) was significantly lower than those in the veridical-image ($M = 86.46\%$) and eyes-closed...
(M = 85.42%) conditions (t_{11} > 3.93, ps < 0.003). For the intermediate torque, the accuracy did not significantly differ among the mirror-image (M = 83.3%), veridical-image (M = 94.8%), and eyes-closed (M = 94.8%) conditions (t_{11} < 1.29, p > 0.22, ns). Similarly, for the strongest torque, the accuracy did not significantly differ among the mirror-image (M = 91.7%), veridical-image (M = 96.9%), and eyes-closed (M = 95.8%) conditions (t_{11} < 1.31, p > 0.21, ns). Thus, the weight distribution information provided by the visual modality was more influential when the torque perceived through the haptic modality was weaker.

We employed a simple linear model to quantify the strength of visual influence.\(^{(1)}\) We denote the level of performance (in terms of proportion correct) in the eyes-closed condition by \(H\) (haptic only), the veridical-image condition by \(HV\) (haptic–visual), and the mirror-image condition by \(HA\) (haptic-reversed, visual). We let \(V\) be the level of performance in a ‘virtual’ pure visual condition in which participants need to visually judge whether the bucket is on the left or right side. Participants are expected to be 100% correct in this trivially easy visual task.

We assume that in the presence of both haptic and visual information, the perceptual decision is made based on a linear summation of information from the two modalities with the visual information weighted by a factor \(w\), and the haptic information weighted by a factor \(1 – w\). Thus, the performance in the visual–haptic condition, \(HV\) or \(HA\) can be expressed as,

\[
HV \text{ or } HA = wV + (1 - w)H.
\]

(1)

Suppose the visual information is completely ignored \((w = 0)\), then the performance is expected to be the same as that in the eyes-closed (haptic only) condition, that is, \(HV = H\). In the other extreme, if the visual information completely dominates \((w = 1)\), then the performance is completely dominated by visual judgments, that is \(HV = V\). Note that \(V = 1\) for the veridical-image trials on which the visual information always indicates correct responses, whereas \(V = 0\) for the mirror-image trials on which the visual information always indicates incorrect responses. In reality, \(w\) is likely to be a value between 0 and 1 as informations from both the visual and haptic modalities are likely to contribute to torque judgments.

In estimating \(w\) (the influence of the visual information), we assume that \(w\) is the same on the veridical-image and mirrored-image trials because those trials were randomly intermixed and participants were unaware of the trial type. We determined \(w\) in the following way. For each weight, we estimated the performance on the cross-modal trials, \(HV\) or \(HA\), based on \(H\) (the level of performance on the eyes-closed trials) and \(V\) (1 on the veridical-image trials and 0 on the mirrored-image trials) using equation (1): \(HV_{\text{estimated}} = w + (1 - w)H\) for the veridical-image trials, and \(HA_{\text{estimated}} = (1 - w)H\) for the mirrored-image trials. To determine the value of \(w\), we analytically found the value of \(w\) that minimized the squared estimation error defined as, \(E = (HV - HV_{\text{estimated}})^2 + (HA - HA_{\text{estimated}})^2\). The estimated \(HV\) and \(HA\) with the optimum values of \(w\) are shown in figure 2a as thin lines.

The value of \(w\) (the linear weight for the visual information) decreased as the haptic torque was increased (mean \(w = 0.385, 0.113,\) and 0.052) \((F_{2.22} = 7.78, p = 0.003)\); \(w\) is significantly higher than zero \((t_{11} = 4.12, p = 0.002)\), for the weakest torque, but is not significantly different from zero for the intermediate or strong torque, \((t_{11} < 1.40, ps > 0.18, ns)\). This analysis confirms that the influence of visual information decreased as haptic information became more reliable.

3 Experiment 2
In experiment 1, participants were not allowed to move the object while making torque judgments, because motion would reveal whether the visual information is veridical or mirrored. We thought that this would trivially reveal the correct direction of the torque.

\(^{(1)}\) We are most grateful to an anonymous reviewer for suggesting this analysis.
Here we tested this assumption. We allowed participants to tilt the object side to side in this experiment, thus providing them with unambiguous information to respond correctly on every trial irrespective of the amount of weight or the visual condition (veridical vs mirrored). Nevertheless, if the visual–haptic interaction demonstrated in experiment 1 is due to relatively automatic cross-modal perceptual interaction, participants may still incorporate the visual information as if it were always veridical, performing better with the veridical image and worse with the mirrored image.

3.1 Method
3.1.1 Participants. All thirteen participants had normal or corrected-to-normal vision, gave informed consent, and were paid for their participation. Data from four of the participants were eliminated from analysis due to chance performance on the haptic-only (eyes-closed) condition even with the strongest torque (suggestive of subnormal haptic perception or deliberate non-compliance). For the remaining nine participants, three of them were male, one of them was left-handed, and two of them were ambidextrous (write with left hand, play sports with right hand).

3.1.2 Procedure. The procedure was identical to that in experiment 1 except that participants were allowed to tilt the object from side to side (within a range of 6° on each side) while making torque direction judgments. Participants typically made their torque judgments within 5 s.

3.2 Results and discussion
Accuracy data for experiment 2 (figure 2b) were submitted to a repeated-measures ANOVA (degrees of freedom, Greenhouse–Geisser corrected when sphericity is violated), with weight and viewing condition as the within-participant factors.

There was a significant main effect of viewing condition ($F_{2,16} = 4.09, p = 0.04, \eta^2 = 0.34$). Accuracy in the mirror-image condition ($M = 82.4\%$) was significantly lower than in the eyes-closed condition ($M = 90.3\%$) ($t_8 = 2.7, p = 0.02$). There was a trend for higher accuracy in the veridical-image condition ($M = 92.6\%$) relative to the mirror-image condition ($M = 82.4\%$) ($t_8 = 1.99, p = 0.07$). Accuracy in the veridical-image condition ($M = 92.6\%$) and the eyes-closed condition ($M = 90.3\%$) did not differ from each other ($t = 0.71, p = 0.49$, ns).

There was also a significant main effect of haptic torque ($F_{2,16} = 21.23, p < 0.001, \eta^2 = 0.73$), with lower performance for the weakest torque ($M = 76.9\%$) compared with the intermediate ($M = 95.4\%$) and strongest ($M = 93.1\%$) torque (all $t_{8} > 4.18, ps < 0.002$). Performance did not differ between the intermediate and strongest torque ($t = 0.9, p = 0.38$, ns).

There was a marginal viewing-condition-by-torque interaction ($F_{2.64,21.10} = 2.65, p = 0.08, \eta^2 = 0.25$). For the weakest torque, the accuracy in the mirror-image condition ($M = 65.3\%$) was significantly lower than those in the veridical-image ($M = 87.5\%$) and eyes-closed ($M = 77.8\%$) conditions ($t_{8} > 2.5, ps < 0.03$). For the intermediate torque, the accuracy did not significantly differ among the mirror-image ($M = 91.7\%$), veridical-image ($M = 95.8\%$), and eyes-closed ($M = 98.6\%$) conditions ($t_{11} < 1.83, p > 0.09$, ns). Similarly, for the strongest torque, the accuracy did not significantly differ among the mirror-image ($M = 90.3\%$), veridical-image ($M = 94.4\%$), and eyes-closed ($M = 94.4\%$) conditions ($t_{11} < 0.65, p > 0.52$, ns). Thus, as in experiment 1, the visual information was more influential when the haptic information was weaker.

We also computed $w$ (the linear weight for the visual information) for each level of torque for each participant. As in experiment 1, the value of $w$ decreased as the haptic torque was increased (mean $w = 0.242, 0.071,$ and $0.049$) ($F_{2,16} = 5.194, p = 0.02$); $w$ was significantly greater than zero ($t_{11} = 3.33, p = 0.007$) for the weakest torque, but was not significantly different from zero for the intermediate and strongest levels of torque ($t_{8} < 1.85, ps > 0.09$).
This analysis confirms that the influence of visual information decreases as haptic information becomes more reliable. The overall value of $w$ (averaged across the three levels of haptic torque) was numerically lower in this experiment (0.12) than in experiment 1 (0.18), but this difference is not statistically significant ($F_{1,19} = 0.64, p = 0.44$). Thus, even when participants were allowed to move the object to be aware of the visual condition (veridical-image vs mirror-image), the visual information still influenced their torque judgments as if it were always veridical.

4 General discussion
When participants held an asymmetrically balanced object, visual information strongly influenced haptic torque judgments. Seeing a mirror-reversed image of the object strongly decreased accuracy, compared to haptic judgment without visual information. Misleading visual information influenced haptic torque judgments, despite the fact that participants were informed that visual information was irrelevant and potentially misleading.

Importantly, the visual effect persisted even when participants were allowed to move the object to know whether the visual image was veridical or mirrored. We predicted that this would make the task trivially easy because the consistency between participants’ intended action and the visual display would reveal the correct response. Surprisingly, the accuracy of haptic torque judgments was still lower when the left–right mirrored visual display was presented than when participants closed their eyes. This suggests that visual information ‘automatically’ influences haptic torque judgments even when participants have the knowledge and incentive to try to ignore the visual information. This inference is consistent with our observation that participants often attempted to look away from the screen in order to focus attention on the haptic sensation in their hand. Our results thus suggest that visual information is cross-modally incorporated into participants’ haptic torque judgments.

A simple assumption that participants linearly weighted the visual and haptic information produced good fits to the data (see the thin lines in figures 2a and 2b). Based on these fits, we determined that the weighting of the visual information in haptic torque judgments decreased as the torque was increased. In other words, visual information influenced haptic torque judgments more strongly when haptic information was weaker. This result is qualitatively consistent with the cue combination model in which it is assumed that the perceptual system combines information from different modalities in a statistically optimal manner by weighting perceptual cues by their sensitivity (reciprocal of variance) (e.g., Ernst and Banks 2002). This model is also consistent with other studies of multisensory discrepancy where the degree of influence by one modality is subject to changes in the type and amount of information received by another modality (e.g., Van Doorn et al. 2010).

An important difference, however, is that in our study the sensitivity to the relevant visual information was virtually 100%, except that it provided correct information only half of the time. Current models of cue combination do not make quantitative predictions about how the weighting of perceptual cues depends jointly on the perceptual sensitivity to the cues and participants’ intentional effort to attend to or ignore the cues. Our results suggest that despite the incentive to ignore the visual information (experiment 1) or to use it correctly when knowing whether the images are veridical or mirrored (experiment 2), the visual information persistently influenced haptic torque judgments as if the perceptual system assumed the visual information was always veridical. At the same time, the visual information was weighted by less than 100% probably because of the explicit knowledge that it was unreliable. Future research should investigate how the explicit knowledge of cue reliability interacts with cue discriminability in determining the influence of a perceptual cue.
The present study complements previous studies that have shown that visual information about volume and dynamics influences perceived weight. When asked to judge the weight of two equally weighted objects, participants typically perceive the one with a bigger volume to be lighter (e.g., Amazeen 1997; Ellis and Lederman 1993; Murray et al. 1999). Streit and colleagues (e.g., Streit et al. 2007a, 2007b) demonstrated that participants judged the weight of a wielded object as lighter if the viewed image of the object appeared to rotate faster in response to a given applied force, and vice versa. The results are consistent with an ‘inertia model’, where the perceived heaviness of a wielded object is a function of the object’s rotational inertia (the object’s resistance to rotational acceleration), and that the perceived inertia is influenced by both haptic and visual information. While these studies investigated how visual information can influence the perceived heaviness of an object, the present study demonstrates a powerful influence of visual information on the perceived weight distribution within an object.

Why might vision exert such a powerful influence on haptic torque judgments? One possibility is that, although the direction of torque is intrinsically a haptic information, there is a consistent association between visual shape imbalance and torque direction. Constant experience of this association throughout the course of development may enable people to use the visual shape imbalance as an efficient cue to estimate the direction of torque, and it may be difficult to ignore visual information even when it is known to be misleading.

Another possibility is that while visual and haptic modalities both provide information about weight distribution, visual information about an object is typically available before haptic information because we tend to see objects before we grasp them. Visual information may therefore tend to be prioritized in order to prepare for appropriate motor responses. This is similar to an expectation theory in weight perception (Ross 1969). But this account has been challenged by findings showing that while visual information presented during haptic exploration can influence the perceived heaviness of an object, information presented only prior to exploration does not influence perceived heaviness. Such results suggest that the visual influence observed in weight perception tasks is a result of sensory integration, as opposed to expectation (Masin and Crestoni 1988). It is possible that when visual and haptic information is available simultaneously, visual information had a strong influence on torque judgments because vision allowed more rapid information extraction and thus operated more efficiently (Heller 1992).

In summary, our results demonstrate that visual information about weight distributions can have a strong influence on haptic torque judgments, even when participants explicitly know that visual information is uninformative. These results suggest that visual information is cross-modally integrated with haptic judgments of weight distribution.

Acknowledgments. This research was supported by a National Institutes of Health grant R01 EY021184. We thank Heeyoung Choo, Brian Levinthal, and Jason Scimeca for helpful comments.

References
Burton G, Turvey M T, 1990 “Perceiving the lengths of rods that are held but not wielded” Ecological Psychology 2 295–324
Ernst M O, Banks M E, 2002 “Humans integrate visual and haptic information in a statistically optimal fashion” Nature 415 429–433
Gibson J J, 1933 “Adaptation, after-effect and contrast in the perception of curved lines” *Journal of Experimental Psychology* 16 1–31
Klein R M, Posner M I, 1974 “Attention to visual and kinesthetic components of skills” *Brain Research* 71 401–411

© 2012 a Pion publication